Theoretical Computer Science 129 (1994) 193–206
Elsevier

Note

Depth of proofs, depth of cut-formulas and complexity of cut formulas

Wenhui Zhang
Economic-Administrative Studies, Telemark College, N-3800 Bø i Telemark, Norway

Communicated by M. Nivat
Received March 1993

Abstract

W. Zhang, Depth of proofs, depth of cut-formulas and complexity of cut formulas, Theoretical Computer Science 129 (1994) 193–206.

The relation between proofs with cuts and proofs without cuts is discussed in this article. The complexity of cut-formulas is redefined to better reflect the structure of cut-formulas which is important to cut-elimination. A cut-elimination strategy based on this definition and an upper bound of cut-elimination are given. Further, it is explained that cut-formulas with depth larger than the depth of the proof can be shortened without increasing the depth of the proof and the complexity of the cut-formulas. A new upper bound based on this fact and the previous upper bound is given.

1. Introduction

This article attempts to find out what the most important part of a cut-formula is, with respect to proof length. A previous study can be found in [3]. Here we provide a modified definition of the complexity of cut-formulas and explain the relation between the complexity of cut-formulas, the depth of cut-formulas and that of cut-free proofs. The main difference is that while the previous study pointed out the importance

Correspondence to: W. Zhang, Economic-Administrative Studies, Telemark Regional College, 3800 Bø i Telemark, Norway.

SSDI 0304-3975(93)E0190-F

of the number of nested quantifiers in a formula, we show in addition that the number of alternating quantifiers is important.

We first present analysis trees [2] and an upper bound for cut-elimination [3] in this section. In Section 2, we present the idea that leads to the modified definition of the complexity of cut-formulas. In Section 3, we present a modified upper bound for cut-elimination. In Section 4, we discuss how unreasonably long cut-formulas can be replaced by shorter formulas. In Section 5, we discuss the relation between the depth and the complexity of cut-formulas and present some simpler upper bounds for cut-elimination.

1.1. Analysis trees

Let φ, ψ be formulas and Δ be a set of formulas, an analysis tree is a proof that uses the following rules:

- Normal rules:

$$A\colon\ \Delta, \varphi, \neg\varphi \quad \text{if } \varphi \text{ is atomic}$$

$$\vee_0\colon\ \frac{\Delta, \varphi}{\Delta, \varphi \vee \psi}, \qquad \vee_1\colon\ \frac{\Delta, \psi}{\Delta, \varphi \vee \psi}$$

$$\wedge\colon\ \frac{\Delta, \varphi \qquad \Delta, \psi}{\Delta, \varphi \wedge \psi}$$

$$\exists\colon\ \frac{\Delta, \varphi(t)}{\Delta, \exists x \mid \varphi(x)}$$

$$\forall\colon\ \frac{\Delta, \varphi(a)}{\Delta, \forall x \mid \varphi(x)} \quad \text{if } a \text{ is not free in } \Delta$$

- Cut rule:

$$\frac{\Delta, \varphi \qquad \Delta, \sim\varphi}{\Delta}$$

Negation $\neg$ applies to atomic formulas only and the symbol $\sim$ is a symbol in the metalanguage which negates the formula following it. We also use $\sim\Delta$ to represent the set of formulas: $\{\sim\psi \mid \psi \in \Delta\}$.

Definition 1.1. The set Π is defined as follows: $\zeta \in \Pi$ iff $\zeta = \sum_{i=0}^{k} a_i \cdot w^i$ with $a_i \geqslant 0$ and $\alpha_k = 1$.

Definition 1.2. Let $\zeta_0 = \sum_{i=0}^{k_0} a_i \cdot w^i \in \Pi$ and $\zeta_1 = \sum_{i=0}^{k_1} b_i \cdot w^i \in \Pi$. The function $\oplus : \Pi \times \Pi \to \Pi$ is defined as follows:

- $\zeta_0 \oplus \zeta_1 \stackrel{\text{def}}{=} \sum_{i=0}^{k_0} \max(a_i, b_i) \cdot w^i$, if $k_0 = k_1$;

- $\zeta_0 \oplus \zeta_1 \stackrel{\text{def}}{=} \zeta_0 \cdot w^{k_1-k_0} \oplus \zeta_1$, if $k_0 < k_1$;
- $\zeta_0 \oplus \zeta_1 \stackrel{\text{def}}{=} \zeta_0 \oplus \zeta_1 \cdot w^{k_0-k_1}$, if $k_0 > k_1$.

This function is associative and commutative. If a and b are numbers, $a \oplus b$ is the same as $\max(a, b)$.

Definition 1.3. The relation $\leqslant$ over Π is defined as follows: $\zeta_0 \leqslant \zeta_1$ iff there exists a $\zeta \in \Pi$ such that $\zeta_0 \oplus \zeta = \zeta_1$.

$(\Pi, \leqslant)$ is a well founded set and $\zeta_0 \leqslant \zeta_1$ implies $\zeta_0 \oplus \zeta \leqslant \zeta_1 \oplus \zeta$.

Definition 1.4. Let $d_1, \ldots, d_k$ with $k \geqslant 0$ be the immediate subproofs of d. The depth of d is defined as follows: $|d| \stackrel{\text{def}}{=} |d_1| \oplus \cdots \oplus |d_k| + 1$.

Definition 1.5. The complexity of a cut-formula is an assignment from the set of formulas to Π and it is define as follows:

- $\rho(\varphi) = \rho(\neg\varphi) \stackrel{\text{def}}{=} w$, if φ is atomic;
- $\rho(\varphi \wedge \psi) = \rho(\varphi \vee \varphi) \stackrel{\text{def}}{=} \rho(\varphi) \oplus \rho(\psi) + 1$;
- $\rho(\forall x \mid \varphi(x)) = \rho(\exists x \mid \varphi(x)) \stackrel{\text{def}}{=} \rho(\varphi) \cdot w$.

Definition 1.6. Let $d_1, \ldots, d_k$ with $k \geqslant 0$ be the immediate subproofs of d. The cut rank is an assignment from the set of analysis trees to Π and the cut rank of d is defined as follows:

- $\rho(d) \stackrel{\text{def}}{=} 1$ if $k = 0$;
- $\rho(d) \stackrel{\text{def}}{=} \rho(d_1) \oplus \cdots \oplus \rho(d_k)$ if the last rule used in d is not a cut;
- $\rho(d) \stackrel{\text{def}}{=} \rho(d_1) \oplus \rho(d_2) \oplus \rho(\psi)$ if ψ is the cut formula of the last rule of d.

Definition 1.7. Let z be either a proof tree or a cut-formula with $\rho(z) = \sum_{i=0}^{k} a_i \cdot w^i$. The functions α and β are defined as follows:

- $\alpha(z) \stackrel{\text{def}}{=} k$;
- $\beta(z) \stackrel{\text{def}}{=} a_0$.

1.2. Cut-elimination

The following is a theorem from a previous paper [3] (the lemmas and the theorem have been reformulated to fit the definitions in Section 1.1 and to ease the comparison with the results of Section 3) which shows the relation between the complexity of the cut-rank of a proof and the upper bound for the proof depth of the corresponding cut-free proof.

Lemma 1.8. (Inversion lemma).
(i) *If* $d \vdash \Delta, \psi_0 \wedge \psi_1$, *we can find* $d_i \vdash \Delta, \psi_i$ $(i=0,1)$ *with* $|d_i| \leqslant |d|$ *and* $\rho(d_i) \leqslant \rho(d)$.
(ii) *If* $d \vdash \Delta, \forall x | \psi(x)$, *we can find* $d' \vdash \Delta, \psi(t)$ *with* $|d'| \leqslant |d|$ *and* $\rho(d') \leqslant \rho(d)$.
(iii) *If* $d \vdash \Delta, \psi_0 \vee \psi_1$, *we can find* $d' \vdash \Delta, \psi_0, \psi_1$ *with* $|d'| \leqslant |d|$ *and* $\rho(d') \leqslant \rho(d)$.

Lemma 1.9. (Reduction lemma).
(i) *If* $d_1 \vdash \Delta, \sim\varphi$ *and* $d_0 \vdash \Delta, \varphi$ *with* $\alpha(d_1), \alpha(d_0) < \alpha(\varphi)$ *and* $\beta(\varphi) = 0$, *we can find* $d \vdash \Delta$ *with* $|d| \leqslant |d_1| + |d_0|$ *and* $\rho(d) \leqslant \rho(d_0) \oplus \rho(d_1) \oplus \rho(\varphi) \cdot w^{-1}$.
(ii) *If* $d_1 \vdash \Delta, \sim\varphi$ *and* $d_0 \vdash \Delta, \varphi$ *with* $(\alpha(d_i) = \alpha(\varphi) \wedge \beta(d_i) = 0) \vee (\alpha(d_i) < \alpha(\varphi))$ *for* $i = 0, 1$, *we can find* $d \vdash \Delta$ *with* $|d| \leqslant |d_1| \oplus |d_0| + 2^{\beta(\varphi)}$ *and* $\rho(d) \leqslant \rho(d_0) \oplus \rho(d_1) \oplus (\rho(\varphi) - \beta(\varphi))$.

Theorem 1.10 (Cut elimination).
(i) *If* $d \vdash \Delta$ *with* $\beta(d) = 0$, *we can find* $d' \vdash \Delta$ *with* $|d'| \leqslant 2^{|d|}$ *and* $\rho(d') \leqslant \rho(d) \cdot w^{-1}$.
(ii) *If* $d \vdash \Delta$, *we can find* $d' \vdash \Delta$ *with* $|d'| \leqslant 2^{\beta(d)} \cdot |d|$ *and* $\rho(d') \leqslant \rho(d) - \beta(d)$.

The cut-elimination in Section 3 follows mainly the strategy indicated by the reduction lemma. The difference is that we try to use the cut-elimination steps as far from the root of the proof-tree as possible.

2. Simple cut-formulas

In this section, we only consider cut-formulas constructed from literals and logical symbols $\forall$ and $\wedge$ and the negation of these (i.e. formulas constructed from literals and logical symbols $\exists$ and $\vee$).

Definition 2.1. Let φ be a formula. $\gamma(\varphi)$ iff
- φ is a literal or
- $\varphi = \varphi_0 \wedge \varphi_1$ and $\gamma(\varphi_0) \wedge \gamma(\varphi_1)$ or
- $\varphi = \forall x | \varphi'(x)$ and $\gamma(\varphi'(x))$.

$\gamma(\varphi)$ iff φ is free of existential quantifier and logical or.

Definition 2.2. We define three notations $\vdash_0$, $\vdash_1$ and δ as follows:
- $d \vdash_0 \Delta$: d is a cut-free proof of Δ.
- $d \vdash_1 \Delta$: d is a proof of Δ and if ψ is a cut formula then either $\gamma(\psi)$ or $\gamma(\sim\varphi)$.
- $\psi \in \delta(\varphi)$ iff

– $\varphi = \psi$ or
– $\varphi = \varphi_0 \wedge \varphi_1$ and $\psi \in \delta(\varphi_0) \cup \delta(\varphi_1)$ or
– $\varphi = \forall x \mid \varphi'(x)$ and $\psi \in \delta(\varphi'(t))$ for any term t.

$\gamma(\psi)$ implies that $\delta(\psi)$ is the set of the subformulas of ψ (with the interpretation: $\varphi(t)$ is a subformula of $\forall x \mid \varphi(x)$ and φ is not a subformula of $\neg\varphi$).

Lemma 2.3 (Inversion lemma). *If $d \vdash_0 \Delta, \psi$ with $\gamma(\psi)$ and if $\psi' \in \delta(\psi)$, we can find $d' \vdash_0 \Delta, \psi'$ with $|d'| \leqslant |d|$.*

This lemma follows from Lemma 1.8 by induction on the structure of ψ.

Lemma 2.4. *If $d_0 \vdash_0 \Delta, \psi$ and $d_1 \vdash_0 \Delta, \Psi$ with $\sim\Psi \subset \delta(\psi), \gamma(\psi)$ and ψ is not invalid, we can find $d \vdash_0 \Delta$ with $|d| \leqslant |d_0| + |d_1|$.*

Proof (by induction on $|d_1|$). (1) The last rule of d_1 is an axiom, i.e. $\psi_0, \neg\psi_0 \in \Delta, \Psi$ where ψ_0 is atomic.
– If $\psi_0, \neg\psi_0 \in \Delta$, let d be Δ.
– If $\psi_0 \in \Delta$ and $\neg\psi_0 \in \Psi$ (the proof is similar if $\neg\psi_0 \in \Delta$ and $\psi_0 \in \Psi$), we can find $d \vdash_0 \Delta, \psi_0$ with $|d| \leqslant |d_0|$ according to the inversion lemma.

(2) The last rule of d_1 is $\forall$, $\vee$ or $\exists$. In this case, there are two possibilities:

– d_1 is $\dfrac{\begin{matrix} d_1^* \\ \Delta', \Psi \end{matrix}}{\Delta, \Psi}$.

By applying the induction hypothesis to d_1^* and d_0, we obtain $d' \vdash_0 \Delta'$ with $|d'| \leqslant |d_0| + |d_1^*|$. Hence we can find $d \vdash_0 \Delta$ with $|d| \leqslant |d'| + 1 \leqslant |d_0| + |d_1|$.

– d_1 is $\dfrac{\begin{matrix} d_1^* \\ \Delta, \Psi' \end{matrix}}{\Delta, \Psi}$.

Since $\sim\Psi' \subset \delta(\psi)$ follows from $\sim\Psi \subset \delta(\psi)$ and $\gamma(\psi)$, by applying the induction hypothesis to d_1^* and d_0, we obtain $d \vdash_0 \Delta$ with $|d| \leqslant |d_0| + |d_1^*| \leqslant |d_0| + |d_1|$.

(3) The last rule of d_1 is $\wedge$. Let d_1 be

$$\dfrac{\begin{matrix} d_{10} \\ \Delta', \Psi \end{matrix} \qquad \begin{matrix} d_{11} \\ \Delta'', \Psi \end{matrix}}{\Delta, \Psi}.$$

By the induction hypothesis, we obtain
- $d'_{10} \vdash_0 \Delta'$ from d_0 and d_{10} with $|d'_{10}| \leqslant |d_0| + |d_{10}|$,
- $d'_{11} \vdash_0 \Delta''$ from d_0 and d_{11} with $|d'_{11}| \leqslant |d_0| + |d_{11}|$.

We obtain $d \vdash_0 \Delta$ from d'_{10} and d'_{11} with $|d| = |d'_{10}| \oplus |d'_{11}| + 1 \leqslant |d_0| + |d_1|$. $\square$

Lemma 2.5. *If $d_0 \vdash_0 \Delta$, ψ and $d_1 \vdash_0 \Delta$, $\sim\psi$ with $\gamma(\psi)$, we can find $d \vdash_0 \Delta$ with $|d| \leqslant 2 \cdot (|d_0| \oplus |d_1|)$.*

Since $\sim\psi \in \delta(\sim\psi)$, this lemma follows from the previous lemma, if ψ is not invalid. If it is invalid, $\delta(\psi)$ contains $\psi_0, \neg\psi_0$ for some atomic ψ_0, according to the inversion lemma, we can find cut-free proofs for Δ, ψ_0 and Δ, $\neg\psi_0$ with depth $|d_0|$. Hence we can find $d \vdash_0 \Delta$ with $|d| \leqslant |d_0| + |d_0|$ according to Lemma 1.9.

Theorem 2.6. *If $d \vdash_1 \Delta$, we can find $d' \vdash_0 \Delta$ with $|d'| \leqslant 2^{|d|}$.*

This theorem follows from Lemma 2.5 by induction on $|d|$. As a consequence, if $d \vdash \Delta$ and the cut formulas in d are pure existential or universal formulas, we can find $d' \vdash_0 \Delta$ with $|d'| \leqslant 2^{|d|}$.

3. General cut-formulas

Combining the idea presented in Section 2 and the general cut-elimination presented in Section 1, we present a modified definition of the complexity of cut-formulas and an upper bound for cut-elimination.

Definition 3.1. We define two notations δ^* and δ':
- $\psi \in \delta^*(\varphi)$ iff
  - $\psi \in \delta(\varphi)$ and
  - ψ is either a conjunction of two formulas or a universally quantified formula.
- $\psi \in \delta'(\varphi)$ iff
  - $\psi \in \delta(\varphi)$,
  - ψ is either a disjunction of two formulas or an existential formula, and
  - all terms appear in ψ also appear in φ.

Both $\delta^*(\varphi)$ and $\delta'(\varphi)$ are subsets of $\delta(\varphi)$. The third condition of the second item guarantees that $\delta'(\varphi)$ is a finite set.

Definition 3.2. The complexity of a cut-formula is re-defined as follows.
- If $\delta'(\varphi)$ is empty, $\rho(\varphi) \stackrel{\text{def}}{=} w$.
- If $\delta'(\varphi) = \{\varphi_i \mid 1 \leqslant i \leqslant n\}$ and there is a φ' such that $\forall x \mid \varphi'(x) \in \delta(\varphi)$ and $\alpha(\varphi') \geqslant \alpha(\varphi_i)$ for $1 \leqslant i \leqslant n$, $\rho(\varphi) \stackrel{\text{def}}{=} (\rho(\varphi_1) \oplus \cdots \oplus \rho(\varphi_n)) \cdot w$.

- If $\varphi=\psi_0\wedge\psi_1$, $\delta'(\varphi)=\{\varphi_i \mid 1\leqslant i\leqslant n\}$ and there is no φ' such that $\forall x|\varphi'(x)\in\delta$ and $\alpha(\varphi')\geqslant\alpha(\varphi_i)$ for $1\leqslant i\leqslant n$, $\rho(\varphi)\stackrel{\text{def}}{=}\rho(\psi_0)\oplus\rho(\psi_1)+1$.
- If $\delta'(\varphi)=\{\varphi\}, \rho(\varphi)\stackrel{\text{def}}{=}\rho(\sim\varphi)$.

In case the cut-formula is a literal, the first item of the definition is applicable. In case the formula is universally quantified, one of the first two items is applicable. In case the formula is a conjunction, one of the first three items is applicable. In case the formula is a disjunction or existentially quantified, we can use the last item. The definition of the cut rank is the same as that in Section 1 with $\rho(\psi)$ being the new complexity of ψ. $\delta'(\varphi)$ is empty if $\gamma(\varphi)$, and $d\vdash_1 \Delta$ is a proof with the cut rank being w.

Lemma 3.3 (Inversion lemma).

(i) *If $d\vdash\Delta,\psi\vee\varphi$ and we can find $d'\vdash\Delta,\psi,\varphi$ with $|d'|\leqslant|d|$ and $\rho(d')\leqslant\rho(d)$.*

(ii) *If $d\vdash\Delta,\psi$ and if ψ' is in $\delta(\psi)$, we can find $d'\vdash\Delta,\psi'$ with $|d'|\leqslant|d|$ and $\rho(d')\leqslant\rho(d)$.*

This lemma is similar to Lemma 1.8. Lemma 1.8 is also true if we use the new definition of the complexity of cut-formulas and the second statement of this lemma follows from the first and the second statement of it (Lemma 1.8 with the new definition of the complexity of cut-formulas).

Lemma 3.4. *If $d_0\vdash\Delta,\psi$ and $d_1\vdash\Delta,\Psi$ with $\rho(\psi)\neq w$, $\sim\Psi\subset\delta^*(\psi)$, $\delta(\psi)\neq\{\psi\}$, $\beta(\psi)=0$ and $\alpha(d_1),\alpha(d_0)<\alpha(\psi)$, we can find $d\vdash\Delta$ with $|d|\leqslant|d_1|+|d_0|$ and $\rho(d)\leqslant\rho(d_0)\oplus\rho(d_1)\oplus\rho(\psi)\cdot w^{-1}$.*

Proof (by induction on $|d_1|$). (1) The last rule of d_1 is an axiom. Δ is also an axiom and we let d be Δ.

(2) The last rule of d_1 is a cut with cut-formula φ. Let d_1 be

$$\frac{\begin{array}{cc} d_{10} & d_{11} \\ \Delta,\varphi,\Psi & \Delta,\sim\varphi,\Psi \end{array}}{\Delta,\Psi}$$

By applying the induction hypothesis to d_0 and d_{10}, we obtain $d'_{10}\vdash\Delta,\varphi$ with $|d'_{10}|\leqslant|d_0|+|d_{10}|$ and $\rho(d'_{10})\leqslant\rho(d_{10})\oplus\rho(d_0)\oplus\rho(\psi)\cdot w^{-1}$.

By applying the induction hypothesis to d_0 and d_{11}, we obtain $d'_{11}\vdash\Delta,\sim\varphi$ with $|d'_{11}|\leqslant|d_0|+|d_{11}|$ and $\rho(d'_{11})\leqslant\rho(d_{11})\oplus\rho(d_0)\oplus\rho(\psi)\cdot w^{-1}$.

By applying the cut rule to d'_{10} and d'_{11}, we obtain $d\vdash\Delta$ with

$$|d|=|d'_{10}|\oplus|d'_{11}|+1\leqslant|d_0|+|d_1|,$$

since $|d_1|=|d_{11}|\oplus|d_{10}|+1$; and

$$\begin{aligned}\rho(d)&=\rho(d'_{11})\oplus\rho(d'_{10})\oplus\rho(\varphi)\\ &\leqslant(\rho(d_{11})\oplus\rho(d_0)\oplus\rho(\psi)\cdot w^{-1})\oplus(\rho(d_{10})\oplus\rho(d_0)\oplus\rho(\psi)\cdot w^{-1})\oplus\rho(\varphi)\\ &=\rho(d_1)\oplus\rho(d_0)\oplus\rho(\psi)\cdot w^{-1},\end{aligned}$$

since $\rho(d_1)=\rho(d_{11})\oplus\rho(d_{10})\oplus\rho(\varphi)$.

(3) The last rule is $\exists$ or $\vee$ with the principal formula φ' in Ψ. Let Ψ be $\Psi' \cup \{\varphi'\}$ (φ' may occur in Ψ') and d_1 be

$$\dfrac{\begin{array}{c} d_1' \\ \Delta, \varphi, \Psi' \end{array}}{\Delta, \varphi', \Psi'}$$

If $\sim\varphi \in \delta^*(\psi)$, we can use induction on d_1' and d_0 to obtain the result. In the following, we assume the opposite. From the conditions in the if-part of this lemma, we obtain $\rho(\varphi) \leqslant \rho(\psi)\cdot w^{-1}$. By applying the induction hypothesis to d_0 and d_1' we obtain $d_1'' \vdash \Delta, \varphi$ with $|d_1''| \leqslant |d_0| + |d_1'|$ and $\rho(d_1'') \leqslant \rho(d_1') \oplus \rho(d_0) \oplus \rho(\psi)\cdot w^{-1}$. By the inversion lemma we obtain $d_0' \vdash \Delta, \sim\varphi$ with $|d_0'| \leqslant |d_0|$ and $\rho(d_0') \leqslant \rho(d_0)$. By applying the cut-rule to d_1'' and d_0', we obtain $d \vdash \Delta$ with

$$|d| = |d_0'| \oplus |d_1''| + 1 \leqslant |d_0| + |d_1'| + 1 = |d_0| + |d_1|;$$

$$\rho(d) \leqslant \rho(d_1'') \oplus \rho(d_0) \oplus \rho(\varphi) \leqslant \rho(d_1') \oplus \rho(d_0) \oplus \rho(\psi)\cdot w^{-1} \oplus \rho(d_0) \oplus \rho(\psi)\cdot w^{-1}$$
$$= \rho(d_1') \oplus \rho(d_0) \oplus \rho(\psi)\cdot w^{-1}.$$

(4) The proof of the other cases (the last rule is $\forall$, $\exists$, $\vee$ or $\wedge$ with principal formula not in Ψ) is similar to the proof of corresponding cases of Lemma 2.4. The difference is that we have to calculate $\rho(d)$ in addition to $|d|$. □

Lemma 3.5. *If $d_1 \vdash \Delta, \sim\psi$ and $d_0 \vdash \Delta, \psi$ with $\alpha(d_1), \alpha(d_0) < \alpha(\psi)$ and $\beta(\psi) = 0$, we can find $d \vdash \Delta$ with $|d| \leqslant 2\cdot(|d_1| \oplus |d_0|)$ and $\rho(d) \leqslant \rho(d_0) \oplus \rho(d_1) \oplus \rho(\psi)\cdot w^{-1}$.*

If $\rho(\psi) = w$ (which implies one of $\gamma(\psi)$ and $\gamma(\sim\psi)$), we refer to Lemma 2.5. Otherwise, either $\delta(\psi) \neq \{\psi\}$ and $\sim(\sim\psi) \in \delta^*(\psi)$ or $\delta(\sim\psi) \neq \{\sim\psi\}$ and $\sim(\psi) \in \delta^*(\sim\psi)$, in both cases, this lemma follows from the previous lemma.

Corollary 3.6. *If $d \vdash \Delta$ with $\beta(d) = 0$, we can find $d' \vdash \Delta$ with $|d'| \leqslant 2^{|d|}$ and $\rho(d') \leqslant \rho(d)\cdot w^{-1}$.*

This corollary follows from Lemma 3.5 by induction on $|d|$.

Lemma 3.7. *If $d_1 \vdash \Delta, \sim\varphi$ and $d_0 \vdash \Delta, \varphi$ with $(\alpha(d_i) = \alpha(\varphi) \wedge \beta(d_i) = 0) \vee (\alpha(d_i) < \alpha(\varphi))$ for $i = 0, 1$, we can find $d \vdash \Delta$ with $|d| \leqslant |d_1| \oplus |d_0| + 2^{\beta(\varphi)}$ and $\rho(d) \leqslant \rho(d_0) \oplus \rho(d_1) \oplus (\rho(\varphi) - \beta(\varphi))$.*

Proof. (by induction on $\beta(\varphi)$). The lemma holds if $\beta(\varphi) = 0$. Assume that $\varphi = \varphi_0 \wedge \varphi_1$ (the proof is similar if $\sim\varphi = \sim\varphi_0 \wedge \sim\varphi_1$). According to the inversion lemma, we can find

- $d_0' \vdash \Delta, \varphi_0$ with $|d_0'| \leqslant |d_0|$ and $\rho(d_0') \leqslant \rho(d_0)$;
- $d_0'' \vdash \Delta, \varphi_1$ with $|d_0''| \leqslant |d_0|$ and $\rho(d_0'') \leqslant \rho(d_0)$;
- $d_1' \vdash \Delta, \sim\varphi_0, \sim\varphi_1$ with $|d_1'| \leqslant |d_1|$ and $\rho(d_1') \leqslant \rho(d_1)$.

Assume that $\alpha(\varphi_0)\geqslant\alpha(\varphi_1)$ (the proof is similar if $\alpha(\varphi_0)\leqslant\alpha(\varphi_1)$). By the induction hypothesis we obtain $d'\vdash\Delta,\ \sim\varphi_1$ from d_1' and d_0' with $|d'|\leqslant|d_1|\oplus|d_0|+2^{\beta(\varphi_0)}$ and $\rho(d')\leqslant\rho(d_1)\oplus\rho(d_0)\oplus(\rho(\varphi_0)-\beta(\varphi_0))$.

(1) In the case that $\alpha(\varphi_1)<\alpha(\varphi_0)$, by applying the cut rule to d' and d_0'' (with φ_1 as the cut-formula), we obtain $d\vdash\Delta$ with

$$\begin{aligned}|d|&\leqslant|d'|\oplus|d_0|+1\leqslant(|d_1|\oplus|d_0|+2^{\beta(\varphi_0)})\oplus|d_0|+1\\&\leqslant|d_1|\oplus|d_0|+2^{\beta(\varphi)}.\end{aligned}$$

since $\beta(\varphi)=\beta(\varphi_0)+1$; and

$$\begin{aligned}\rho(d)&=\rho(d')\oplus\rho(d_0)\oplus\rho(\varphi_1)\leqslant(\rho(d_1)\oplus\rho(d_0)\oplus(\rho(\varphi_0)\\&\quad-\beta(\varphi_0)))\oplus\rho(d_0)\oplus\rho(\varphi_1)\\&=\rho(d_1)\oplus\rho(d_0)\oplus(\rho(\varphi)-\beta(\varphi)),\end{aligned}$$

since $\rho(\varphi)-\beta(\varphi)=(\rho(\varphi_0)-\beta(\varphi_0))\oplus\rho(\varphi_1)$.

(2) In the case that $\alpha(\varphi_0)=\alpha(\varphi_1)$, by applying the induction hypothesis to d' and d_0'', we obtain $d\vdash\Delta$ with

$$\begin{aligned}|d|&\leqslant|d'|\oplus|d_0|+2^{\beta(\varphi_1)}\leqslant(|d_1|\oplus|d_0|+2^{\beta(\varphi_0)})\oplus|d_0|+2^{\beta(\varphi_1)}\\&\leqslant|d_1|\oplus|d_0|+2^{\beta(\varphi)},\end{aligned}$$

since $\beta(\varphi)=\beta(\varphi_0)\oplus\beta(\varphi_1)+1$; and

$$\begin{aligned}\rho(d)&\leqslant\rho(d')\oplus\rho(d_0)\oplus(\rho(\varphi_1)-\beta(\varphi_1))\\&\leqslant(\rho(d_1)\oplus\rho(d_0)\oplus(\rho(\varphi_0)-\beta(\varphi_0)))\oplus\rho(d_0)\oplus(\rho(\varphi_1)-\beta(\varphi_1))\\&=\rho(d_1)\oplus\rho(d_0)\oplus(\rho(\varphi)-\beta(\varphi)),\end{aligned}$$

since $\rho(\varphi)-\beta(\varphi)=(\rho(\varphi_0)-\beta(\varphi_0))\oplus(\rho(\varphi_1)-\beta(\varphi_1))$. □

Corollary 3.8. *If $d\vdash\Delta$, we can find $d'\vdash\Delta$ with $|d'|\leqslant2^{\beta(d)}\cdot|d|$ and $\rho(d')\leqslant\rho(d)-\beta(d)$.*

This corollary follows from Lemma 3.7 by induction on $|d|$.

Theorem 3.9. *If $d\vdash\Delta$ with $\rho(d)=\sum_{i=0}^{k}a_i\cdot w^i$ and $a_k>0$, we can find $d'\vdash\Delta$ with $|d'|\leqslant2^{2^{a_0}\cdot|d|}$ and $\rho(d')\leqslant\sum_{i=1}^{k}a_i\cdot w^{i-1}$.*

This theorem follows from Corollary 3.8 and Corollary 3.6.

Corollary 3.10. *If $d\vdash\Delta$ we can find a cut-free $d'\vdash\Delta$ with $|d'|\leqslant s(\rho(d),|d|)$ where $s(1,n)=n$ and $s(\zeta\cdot w+a,n)=s(\zeta,2^{2^{a}\cdot n})$.*

4. Depth of cut-formulas

The object is to show the relation between the depth of cut-formulas and the depth of proof. We prove that cut-formulas with depth larger than the depth of the proof-tree can be replaced by shorter cut-formulas. The result is similar to Lemma 3.3 of [1]. We make it clear that the complexity (by Definition 3.2) of cut-formulas does not increase and make the bound of the depth of cut-formulas more accurate. In addition, we also indicate how the shorter cut-formulas are constructed and it may be useful for proof presentation.

Definition 4.1. The depth of a formula $|\psi|$ is defined as follows:

- $|\psi| = |\neg\psi| = 1$;
- $|\psi_0 \wedge \psi_1| = |\psi_0 \vee \psi_1| = |\psi_0| \oplus |\psi_1| + 1$;
- $|\forall x|\psi(x)| = |\exists x|\psi(x)| = |\psi(x)| + 1$.

$|\psi|$ is the original complexity (as in [2]) of the formula ψ.

Definition 4.2. ψ is a subformula of φ at level k, written as $l(\psi, \varphi, k)$ is defined as follows:

- $l(\varphi, \varphi, 0)$;
- $l(\psi, \forall x|\varphi(x), k)$ iff $l(\psi, \varphi(x), k-1)$;
- $l(\psi, \exists x|\varphi(x), k)$ iff $l(\psi, \varphi(x), k-1)$;
- $l(\psi, \varphi_0 \wedge \varphi_1, k)$ iff $l(\psi, \varphi_0, k-1)$ or $l(\psi, \varphi_1, k-1)$;
- $l(\psi, \varphi_0 \vee \varphi_1, k)$ iff $l(\psi, \varphi_0, k-1)$ or $l(\psi, \varphi_1, k-1)$;

Note that the k in $l(\psi, \varphi, k)$ is not necessary unique, since ψ may occur in many subformulas of φ. Let $\varphi[p/\psi, k]$ be the result of substituting p (in the following, the symbols p and p_i are used to represent literals) for all ψ at level k of φ (if ψ appears at other levels of φ, it will not be changed). Let $\varphi[x/t]$ be the result of substituting x for t in φ. The following are some properties of the substitution.

- $\varphi[p/\psi, 0] =$ (if $\varphi \neq \psi$ then φ else p).
- $\varphi[p/\psi, k] = \varphi$, if φ is a literal and $k \geqslant 1$.
- $(\forall x|\varphi(x))[p/\psi(x), k] = \forall x|(\varphi(t)[p/\psi(t), k-1][x/t])$ for $k \geqslant 1$.
- $(\exists x|\varphi(x))[p/\psi(x), k] = \exists x|(\varphi(t)[p/\psi(t), k-1][x/t])$ for $k \geqslant 1$.
- $(\varphi_0 \wedge \varphi_1)[p/\psi, k] = \varphi_0[p/\psi, k-1] \wedge \varphi_1[p/\psi, k-1]$ for $k \geqslant 1$.
- $(\varphi_0 \vee \varphi_1)[p/\psi, k] = \varphi_0[p/\psi, k-1] \vee \varphi_1[p/\psi, k-1]$ for $k \geqslant 1$.
- $\sim\varphi[p/\psi, k] = (\sim\varphi)[\neg p/\sim\psi, k]$ for $k \geqslant 0$.
- $\rho(\varphi[p/\psi, k]) \leqslant \rho(\varphi)$ for $k \geqslant 0$.

The last two items are needed in the proof of Theorem 4.5 and the rest is needed in the proof of Lemma 4.3.

Lemma 4.3. *If $d \vdash \Delta, \varphi$ with $|d| \leqslant k$, we can find $d' \vdash \Delta, \varphi[p/\psi, k]$ with $|d'| \leqslant |d|$ and $\rho(d') \leqslant \rho(d)$.*

Proof (by induction on d). (1) $|d|=1$.
If φ is a literal, $\varphi[p/\psi,k]=\varphi$.
If φ is not a literal, Δ is an axiom.

(2) φ is not the principal formula of the last rule of d.
The last rule of d is $\wedge$ or a cut. Let d be

$$\frac{\begin{array}{cc} d_1 & d_0 \\ \Delta', \varphi & \Delta'', \varphi \end{array}}{\Delta, \varphi}$$

By the induction hypothesis, we can find

$$d_1' \vdash \Delta', \varphi[p/\psi,k] \text{ with } |d_1'| \leqslant |d_1| \text{ and } \rho(d_1') \leqslant \rho(d_1), \text{ and}$$

$$d_0' \vdash \Delta'', \varphi[p/\psi,k] \text{ with } |d_0'| \leqslant |d_0| \text{ and } \rho(d_0') \leqslant \rho(d_0).$$

By combining d_0' and d_1', we obtain $d' \vdash \Delta, \varphi[p/\psi,k]$ with $|d'| \leqslant |d|$. In the case that the rule is $\wedge$, we obtain

$$\rho(d') = \rho(d_1') \oplus \rho(d_0') \leqslant \rho(d_1) \oplus \rho(d_0) = \rho(d).$$

In the case that the rule is a cut with φ_0 as the cut formula, we obtain

$$\rho(d') = \rho(d_1') \oplus \rho(d_0') \oplus \varphi_0 \leqslant \rho(d_1) \oplus \rho(d_0) \oplus \varphi_0 = \rho(d).$$

The last rule is $\forall$, $\exists$ or $\vee$.
The proof is similar to the previous case.

(3) φ is the principal formula of the last rule of d.
The last rule of d is $\exists$. Let φ be $\exists x \mid \varphi_0(x)$. Assume d is (the case where $\exists x \mid \varphi_0(x)$ not in the conclusion of d_1 being similar):

$$\frac{\begin{array}{c} d_1 \\ \Delta, \exists x \mid \varphi_0(x), \varphi_0(t) \end{array}}{\Delta, \exists x \mid \varphi_0(x)}$$

By the induction hypothesis, we can find

$$d_1' \vdash \Delta, (\exists x \mid \varphi_0(x))[p/\psi(x),k], \varphi_0(t) \text{ with } |d_1'| \leqslant |d_1| \text{and } \rho(d_1') \leqslant \rho(d_1).$$

By the induction hypothesis, we can find

$$d_1'' \vdash \Delta, (\exists x \mid \varphi_0(x))[p/\psi(x),k], \varphi_0(t)[p/\psi(t),k-1] \text{ with}$$

$$|d_1''| \leqslant |d_1'| \text{and } \rho(d_1'') \leqslant \rho(d_1').$$

Hence we can find

$$d' \vdash \Delta, (\exists x \mid \varphi_0(x))[p/\psi(x),k], \text{ with}$$

$$|d'| \leqslant |d_1''| + 1 \leqslant |d| \text{ and } \rho(d') \leqslant \rho(d_1'') \leqslant \rho(d).$$

The last rule is $\forall$, $\vee$ or $\wedge$.
The proof is similar to the previous case. $\square$

Lemma 4.4. *If $d \vdash \Delta, \varphi$ with $|d| \leqslant k$ and there are $n+1$ formulas $\varphi_0, \ldots, \varphi_n$ with the property $l(\varphi_i, \varphi, k)$, we can find $d' \vdash \Delta, \varphi[p_0/\varphi_0, k] \cdots [p_n/\varphi_n, k]$ with $|d'| \leqslant |d|$, $\rho(d') \leqslant \rho(d)$.*

This lemma follows from the previous lemma by induction on the number of different subformulas at level k of φ.

Theorem 4.5. *If the last rule of $d \vdash \Delta$ is a cut with φ as the cut-formula, we can find $d' \vdash \Delta$ such that d' is the result of replacing φ in d by φ' with $|\varphi'| \leqslant |d|$ and $\rho(\varphi') \leqslant \rho(\varphi)$.*

Proof. Assume that $d_0 \vdash \Delta, \varphi$ and $d_1 \vdash \Delta, \sim\varphi$. Let k be $|d_1| \oplus |d_0|$ and $\varphi_0, \ldots, \varphi_n$ be the only formulas with the property $l(\varphi_i, \varphi, k)$. We obtain that $\sim\varphi_0, \ldots, \sim\varphi_n$ are the only formulas with the property $l(\sim\varphi_i, \sim\varphi, k)$. According to the previous lemma, we can find

- $d_0' \vdash \Delta, \varphi[p_0/\varphi_0, k] \cdots [p_n/\varphi_n, k]$ with $|d_0'| \leqslant |d_0|$ and $\rho(d_0') \leqslant \rho(d_0)$
- $d_1' \vdash \Delta, (\sim\varphi)[\neg p_0/\sim\varphi_0, k] \cdots [\neg p_n/\sim\varphi_n, k]$ with $|d_1'| \leqslant |d_1|$ and $\rho(d_1') \leqslant \rho(d_1)$

Since $(\sim\varphi)[\neg p_0/\sim\varphi_0, k] \cdots [\neg p_n/\sim\varphi_n, k] = \sim\varphi[p_0/\varphi_0, k] \cdots [p_n/\varphi_n, k]$, we let φ' be $\varphi[p_0/\varphi_0, k] \cdots [p_n/\varphi_n, k]$. We obtain $|\varphi'| \leqslant k+1$ and $\rho(\varphi') \leqslant \rho(\varphi)$. □

Corollary 4.6. *If $d \vdash \Delta$, we can find $d' \vdash \Delta$ with $|d'| \leqslant |d|$, $\rho(d') \leqslant \rho(d)$ and that if ψ is a cut-formula in d', then $|\psi| \leqslant |d|$.*

Proof. According to Theorem 4.5, we can replace all cut-formulas with depth $\geqslant |d|$ with formulas with depth $|d|$ without increasing the proof depth and the complexity of the cut-formulas. □

Corollary 4.7. *If $d \vdash \Delta$ we can find a cut free $d' \vdash \Delta$ with $|d'| \leqslant 2_{|d|}^{|d|}$.*

This corollary follows from Corollary 4.6 and the cut-elimination theorem in [2].

5. Depth and complexity

We discuss the relation between $|\psi|$ and $\rho(\psi)$ and apply this relation to the theorems in the previous sections to obtain simpler upper bounds for cut-elimination.

Lemma 5.1. *If $\rho(\psi) = \sum_{i=0}^{k} a_i \cdot w^i$, then $k \leqslant |\psi|$ and $a_i \leqslant |\psi| - k + i$ for $i = 0, \ldots, k$.*

Proof. It is obvious that $k \leqslant |\psi|$ and the rest of this lemma is proved by induction on k and a_0 as follows.

(1) If $k=1$, we have $a_0 \leqslant |\psi| - 1$ and $a_1 = 1 \leqslant |\psi| - k + 1$.

(2) If $\rho(\psi)=(\rho(\psi_1)\oplus\cdots\oplus\rho(\psi_n))\cdot w$, there is a $\psi_m (1\leqslant m\leqslant n)$ with $\rho(\psi_m)=\sum_{j=0}^{k_0} b_j\cdot w^j$ such that $a_i=b_{j0}$ and $i=j_0+k-k_0-1$ for some j_0. According to the induction hypothesis, we have

$$a_i=b_{j0}\leqslant|\psi_m|-k_0+j_0\leqslant|\psi|-1-k_0+j_0\leqslant|\psi|-k+i.$$

(3) If $\rho(\psi)=\rho(\psi_0)\oplus\rho(\psi_1)+1$, we have two cases:
In the case that $i=0$: There is a ψ_m $(0\leqslant m\leqslant 1)$ with $\rho(\psi_m)=\sum_{j=0}^{k_0} b_j\cdot w^j$ such that $a_0=b_0+1$ and $k=k_0$. According to the induction hypothesis, we have

$$a_0=b_0+1\leqslant(|\psi_m|-k_0)+1\leqslant|\psi|-k.$$

In the case that $i\geqslant 1$: There is a $\psi_m(0\leqslant m\leqslant 1)$ with $\rho(\psi_m)=\sum_{j=0}^{k_0} b_j\cdot w^j$ such that $a_i=b_{j0}$ and $i=j_0+k-k_0$ for some j_0. According to the induction hypothesis, we have

$$a_i=b_{j0}\leqslant|\psi_m|-k_0+j_0\leqslant(|\psi|-1)-k+i\leqslant|\psi|-k+i. \qquad \square$$

Lemma 5.2. *If $\rho(\psi_0)\oplus\rho(\psi_1)=\sum_{i=0}^{k_0} a_i\cdot w^i$, then $a_i\leqslant|\psi_0|\oplus|\psi_1|-k+i$.*

Proof. Let $\rho(\psi_0)=\sum_{i=0}^{k_0} a_i'\cdot w^i$ and $\rho(\psi_1)=\sum_{i=0}^{k_1} a_i''\cdot w^i$. According to the previous lemma, we obtain: $a_i'\leqslant|\psi_0|-k_0+i$ and $a_i''\leqslant|\psi_1|-k_1+i$. Assume $k=k_0$ (the case $k=k_1$ being similar).

- if $i-k+k_1\geqslant 0$, then

$$\begin{aligned} a_i=a_i'\oplus a_{i-k+k_1}''&\leqslant(|\psi_0|-k+i)\oplus(|\psi_1|-k_1+i-k+k_1)\\ &=(|\psi_0|-k+i)\oplus(|\psi_1|+i-k)=|\psi_0|\oplus|\psi_1|-k+i.\end{aligned}$$

- if $i-k+k_1<0$, then

$$a_i=a_i'\leqslant|\psi_0|-k+i\leqslant|\psi_0|\oplus|\psi_1|-k+i. \qquad \square$$

Lemma 5.3. *If $d\vdash\Delta$, we can find $d'\vdash\Delta$ with $|d'|\leqslant|d|$, $\rho(d')\leqslant\rho(d)$ and $a_i\leqslant|d|-k+i$ if $\rho(d')=\sum_{i=0}^{k} a_i\cdot w^i$.*

Proof. According to Corollary 4.6, we can find d' with $\rho(d')\leqslant\rho(d)$ and if ψ_i for $i=1,\ldots,n$ are all of the cut formulas in d', $\rho(\psi_1)\oplus\cdots\oplus\rho(\psi_n)\leqslant|d|$. Hence, $a_i\leqslant|d|-k+i$ according to Lemma 5.2.

Theorem 5.4. *If $d\vdash\Delta$ with $\rho(d)=\sum_{i=0}^{k} a_i\cdot w^i$, we can find $d'\vdash\Delta$ with $|d'|\leqslant|d|$ and $\rho(d')\leqslant\sum_{i=0}^{k'}\min(a_{i-k'+k},|d|-k'+i)\cdot w^i$ where $k'=\min(k,|d|)$.*

Proof. Let $\rho(d')=\sum_{i=0}^{k_0} b_i\cdot w^i$. According to the previous lemma, we have $b_i\leqslant a_{i-k+k_0}$, $b_i\leqslant|d|-k_0+i$ and $k_0\leqslant k'$. Hence

$$\begin{aligned}\rho(d')&\leqslant\textstyle\sum_{i=0}^{k_0}\min(a_{i-k+k_0},|d|-k_0+i)\cdot w^i\\ &\leqslant w^{k'-k_0}\cdot\textstyle\sum_{i=0}^{k_0}\min(a_{i-k+k_0},|d|-k_0+i)\cdot w^i\\ &\leqslant\textstyle\sum_{i=0}^{k'}\min(a_{i-k'+k},|d|-k'+i)\cdot w^i.\end{aligned}$$

Using k' instead of k in this theorem makes it look complicate. The main reason is that we want to make sure that $|d|-k'+i \geqslant 0$ in the theorem. $\square$

Corollary 5.5. *If* $d \vdash \Delta$ *with* $\rho(d') = \sum_{i=0}^{k} a_i \cdot w^i$, *we can find a cut free* $d' \vdash \Delta$ *with* $|d'| \leqslant s(\zeta, |d|)$ *where* $\zeta = \sum_{i=0}^{k'} \min(a_{i-k'+k}, |d|-k'+i) \cdot w^i$ *with* $k' = \min(k, |d|)$.

Corollary 5.6. *If* $d \vdash \Delta$, *we can find a cut free* $d' \vdash \Delta$ *with* $|d'| \leqslant 2_{\alpha(d)}^{2^{\beta(d)+1} \cdot |d|}$.

Corollary 5.7. *If* $d \vdash \Delta$, *we can find a cut free* $d' \vdash \Delta$ *with* $|d'| \leqslant 2_{\alpha(d)+1}^{2 \cdot |d|}$.

6. Summary

Theorem 4.5 shows that we can replace cut-formulas in a proof by formulas with depth less than or equal to the depth of the proof without changing the structure of the proof. As a consequence of this theorem and theorem 3.9, four upper bounds of cut-elimination are given in respectively Corollaries 5.5–5.7 and Corollary 4.7. The simplest upper bound is that in Corollary 4.7. Usually, $\alpha(d)$ is much less than $|d|$ and hence $2_{\alpha(d)+1}^{2 \cdot |d|}$ is a better upper bound. If we also know the value of $\beta(d)$, we can use $2_{\alpha(d)}^{2^{\beta(d)+1} \cdot |d|}$ as an upper bound. The upper bound in Corollary 5.5 makes use of all details of a cut-rank.

Roughly speaking, $\alpha(\psi)$ is the number of the times (possibly plus one for the atomic formulas in ψ) that the quantifiers in some path of ψ (represented as a tree) shifted from $\forall$ to $\exists$ or from $\forall$ to $\vee$ to $\forall$ and correspondingly from $\exists$ to $\forall$ or from $\exists$ to $\wedge$ to $\exists$. We may say that it is the primary complexity of ψ. $\beta(\psi)$ is the number of $\wedge$ and $\vee$ which have to be removed before one obtains all quantified formulas with the same primary complexity as that of ψ. The combination of Definitions 3.2, 1.6 and the upper bounds shows that the number of alternating quantifiers in a cut-formula is very important with respect to the possibility of achieving short proofs. It provides a better theoretical basis (compared with the previous definitions of the complexity of cut-formulas and cut-elimination theorems [2, 3]) for constructing good cut-formulas in first-order logic.

The main subject of this article is to clarify what the most important part of a cut formula is, the number of cut-formulas does not matter in the discussion and the relation between the number of nodes and cut-formulas is only indirectly discussed through the discussion of proof depth. Further study is needed to establish the relation between the number of nodes, the number of cut-formulas and the complexity of cut-formulas in a proof.

References

[1] J. Krajicek and P. Pudlak, The number of proof lines and the size of proofs in first order logic, *Arch. Math. Logic* **27** (1988) 69–84.

[2] H. Schwichtenberg, Proof theory: some applications of cut-elimination, in: J. Barwise, ed., *Handbook of Mathematical Logic* (North-Holland, Amsterdam, 1977).

[3] W. Zhang, Cut elimination and automatic proof procedures, *Theoret. Comput. Sci.* **91** (1991) 265–284.